Scientific Research and Essays Vol. 5(17), pp. 2480-2490, 4 September, 2010
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 ©2010 Academic Journals

Full Length Research Paper

Increasing the reliability of skin detectors

Alaa Y. Taqa[1]* and Hamid A. Jalab[2]

Department of Computer System and Technology, University of Malaya, 50603, Kuala Lumpur, Malaysia.

Accepted 13 July, 2010

Skin detection is a well-known image processing technique that has been used in many applications such as video surveillance, naked image filters, and face detection. This paper proposes a reliable skin detection method that integrates both color and texture features. To increase the reliability of the skin detection process, neighborhood pixel information is incorporated into the proposed method. Texture features were estimated using statistical measures as range, standard deviation, and entropy. Back propagation artificial neural network is then used to learn features and classify any given input. In this work, three skin detectors based on pre-defined rules of skin color tones, texture features only, and a combination of both color and texture features have been constructed. Furthermore, the paper analyzes and compares the obtained results from the proposed skin detector with pre-defined color tones and texture feature-based skin detectors to show the level of robustness of the proposed skin detector. It found that the proposed skin detection method achieved a true positive rate of approximately 95.6176% and a false positive rate of approximately 0.8795%. Experimental results showed that our approach is more efficient compared with other state-of-the-art color-based or texture-based skin detector approaches.

Key words: Skin detector, supervised learning, back propagation ANN, Image texture features, Image segmentation.

INTRODUCTION

Skin detection is one of the important techniques in image processing and the most distinctive and widely used key to many applications such as face detection (Kovac, 2003) , face tracking (Dadgostar et al., 2005), human motion analysis (Gavrila, 1999), naked images filters (Fleck et al., 1996) and others. Skin detection is used to determine the image pixels related to human skin. Color is a useful cue to extract skin pixels. One of the major issues in using skin color in skin detection is how to choose a suitable color space. Numerous color models are used today because color science is a broad field encompassing many areas of applications. The most common color space models are RGB, CMY, and CMYK (Gonzalez and Woods, 2002); Hue, Saturation, and Intensity (HIS) (Umbaugh 1998; Singh et al., 2003); Hue, Saturation, and Value (HSV) (Lin et al., 2003); Normalized RGB (Chan et al., 1999; Vezhnevets et al., 2003); and YCbCr (Umbaugh, 1998; Lin et al., 2003). Many skin models have been developed based on RGB (Vezhnevets et al., 2003), but these approaches are not robust enough to handle different lighting conditions and complex backgrounds containing surfaces and objects with skin-like colors. Many researchers (Gasparini et al., 2005; Taqa and Al-Sulaifanie, 2010) have used pixel-based algorithms as main methods for skin detection. Nevertheless, few skin detection methods have been constructed based on a pixel and its neighbors (Ruiz-del-Solar and Verschae, 2004).

Some researchers have used traditional techniques (Jones and Rehg, 2000; Ghouzali et al., 2008; Maskooki et al., 2009), while others used intelligence (Brown et al., 2001; Ahmed et al., 2007; Subramaniani et al., 2008) to detect skin pixels.

This paper proposes a reliable skin detection method using an artificial neural network (ANNSD) that integrates both color and texture features involving the information of a pixel and its neighbors. To determine the decision

*Corresponding author. E-mail: alaa_taqa@um.edu.my, hamidjalab@um.edu.my.

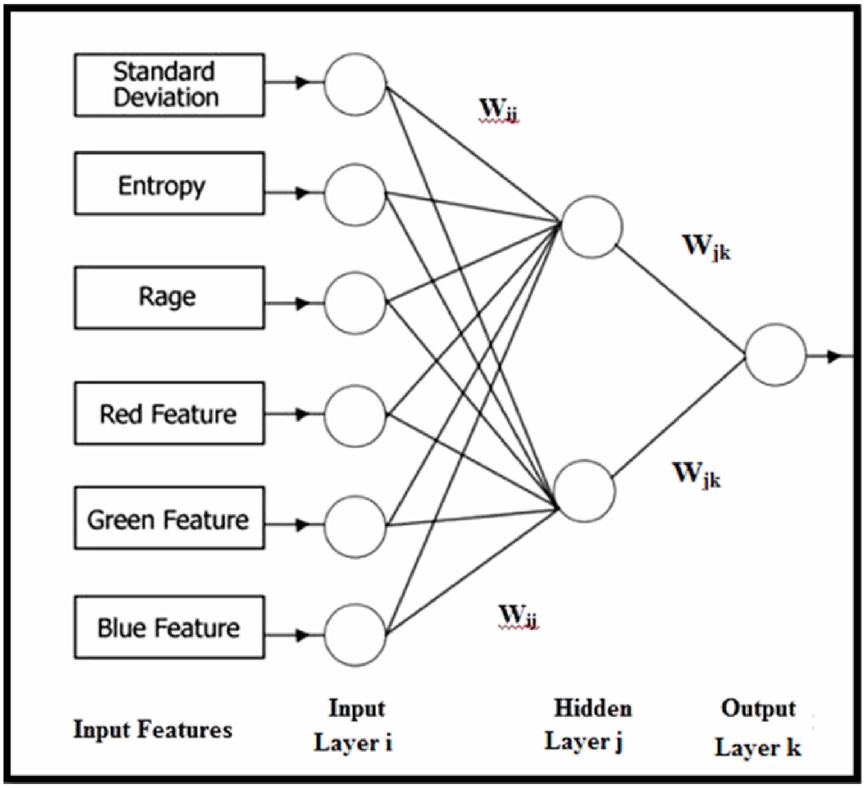

Figure 1. The used back propagation ANN topology.

rule of these features, a back propagation artificial neural network (ANN) is used. To select the significant contribution of increasing the accuracy of the skin detector, three different types of skin detectors are constructed in this paper: Those using colors based on pre-defined rules, skin detector, and a combination of texture features with color feature-based skin detector. This will aid in constructing robust and highly reliable skin detectors. The rest of the paper is organized as follows: Section 2 presents the materials and methods used in the proposed skin detection; Section 3 describes the results obtained from the three detectors; and Section 4 presents the evaluations of each detector. The last section contains the conclusion and suggestions for future work.

MATERIALS AND METHODS

Classification has been one of the most active areas of neural networks since the late 1980s. Neural networks have been successfully applied to a variety of real world classification tasks in industry, business, and science (Sivanandam et al., 2006). The proposed skin detection method -based on ANN- combines both color and texture features. To increase the reliability of the skin detection process, neighborhood pixel information is incorporated into the proposed method.

The color features are extracted directly from the pixels, and the texture features of the scanned windows over the image are extracted using a statistical approach and then feature vector is produced. To determine the decision rule of these features, a back-propagation ANN (Schalkoff, 1999; Taysi, 2010) is used. The architecture of the used back propagation ANN is shown in Figure 1. It consists of: One input layer, one hidden layer and one output layer. There are six neurons in the input layer, each one representing one feature of the feature vector; two neurons in the hidden layer; and one in the output layer, which equals to 1, when the input features represent the skin region, and -1, when the input features represent non-skin regions.

Throughout feature extraction steps, two windows are moved over an image; the size of the first window is 3 × 3 and that of the second is 9 × 9. The color features (Red, Green, and Blue) of the centered pixel of the first window are extracted. The texture features are extracted using a statistical approach, in which different properties are computed using standard statistical measures as standard deviation, range, and entropy. The first two static features (standard deviation and range) are estimated from pixels within the first window, whereas, entropy is estimated from the pixels within the second window. These six features are used as input to the neural network. Figure 2 shows the block diagram of the proposed skin detection method.

The proposed skin detector algorithm

The proposed ANNSD (Artificial Neural Network Skin Detection) consists of the following phases:

Creation of skin and non-skin image database

This stage involves collecting samples of different human skin-colored pixels from a variety of people under different illumination conditions (skin pixels without background), as well as a variety of

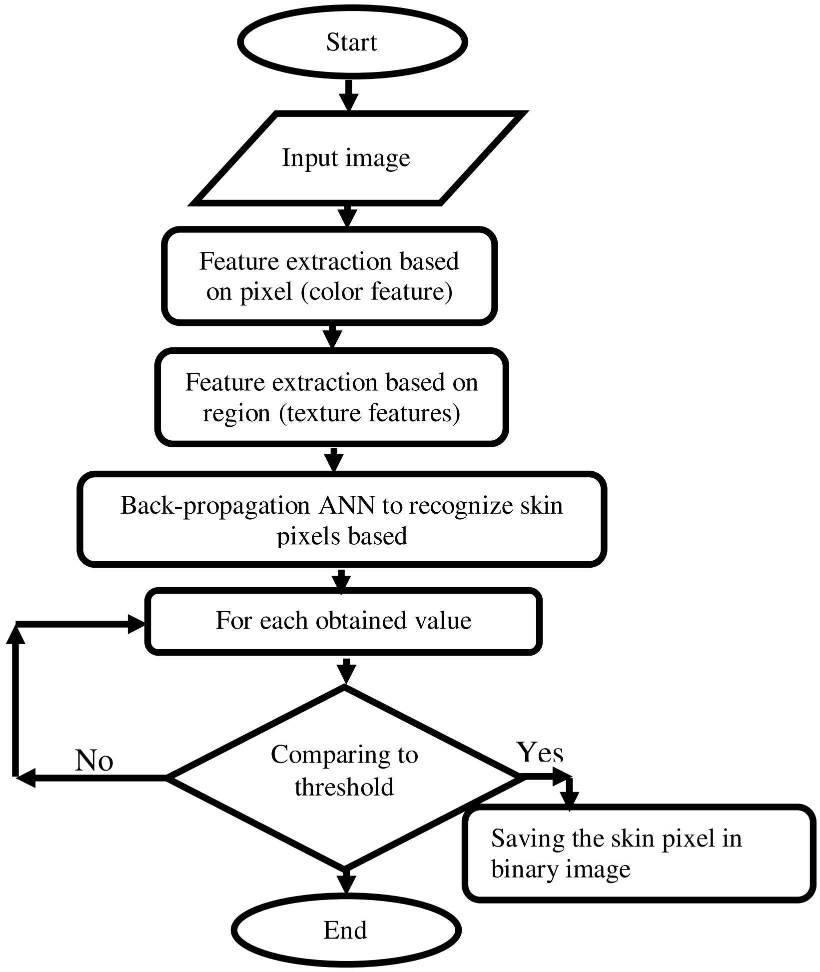

Figure 2. The proposed skin detection method.

non-skin colored pixels. The image is examined manually to determine whether it contains skin. If no skin is present, the image is placed in the non-skin group. In the skin image group, regions of skin pixels are manually extracted using Adobe Photoshop. In labeling skin, an attempt is done to exclude the eyes, hair, clothes, mouth opening, and lips. The collected data are divided into three subsets: Training, validation, and testing. The training set is used as the primary set of data applied to the ANN for learning and adaptation, with 351.228 skin and 428.602 non-skin pixels manually segmented from 87 images. The validation set is used to further refine the ANN construction and is considered an important guard. The test set includes different images with simple and complex backgrounds, indoor and outdoor settings, and different image sizes and skin colors used to measure the performance of the ANN. It has 1,119.129 different pixel types manually segmented from 95 images.

ANN training phase

There are two main stages in this phase: Extracting features, and training the detector to learn these features. The input image may contain objects other than skin; thus, the detector is trained to recognize skin and non-skin features. Throughout the feature extraction stage, three texture features are estimated using a statistical approach, computing different properties using three statistical measures: Standard deviation, entropy and maximum-minimum range. Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image (Gonzalez and Woods, 2002). Entropy is defined as:

$$\mathbf{E_r} = \sum_{i=1}^{n} \mathbf{P_r}(\mathbf{x_i}) * \log_2 (\mathbf{P_r}(\mathbf{x_i})) \tag{1}$$

$$\mathbf{E_g} = \sum_{i=1}^{n} \mathbf{P_g}(\mathbf{x_i}) * \log_2 (\mathbf{P_g}(\mathbf{x_i})) \tag{2}$$

$$\mathbf{E_b} = \sum_{i=1}^{n} \mathbf{P_b}(\mathbf{x_i}) * \log_2 (\mathbf{P_b}(\mathbf{x_i})) \tag{3}$$

The average entropy matrix will be:

$$E = \frac{\mathbf{E_r} + \mathbf{E_g} + \mathbf{E_b}}{3} \tag{4}$$

All statistical measures are computed for multi-channel image

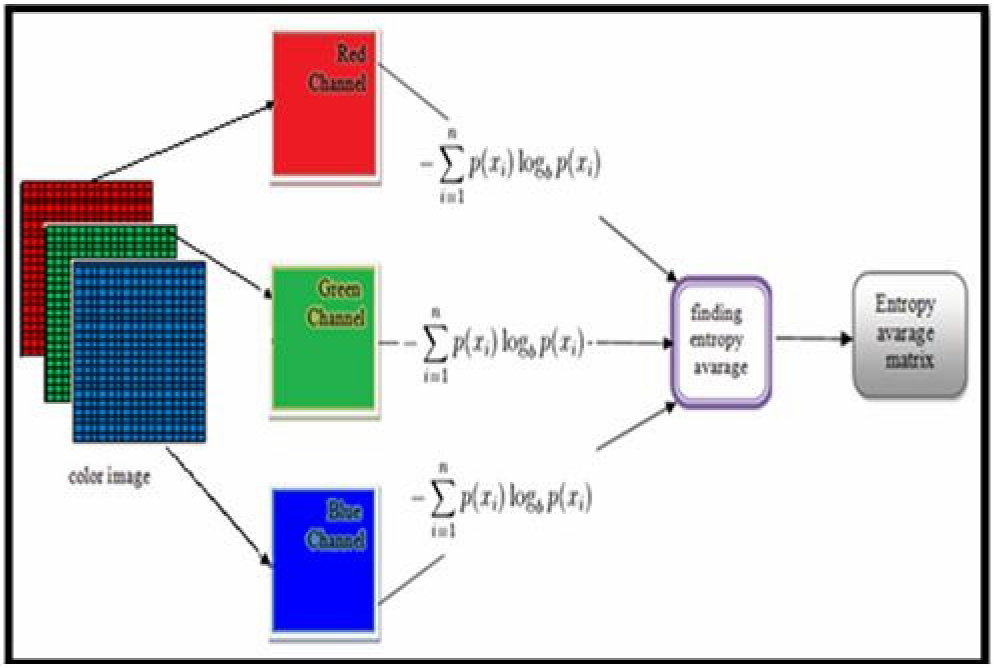

Figure 3. Computing scheme entropy for colored image.

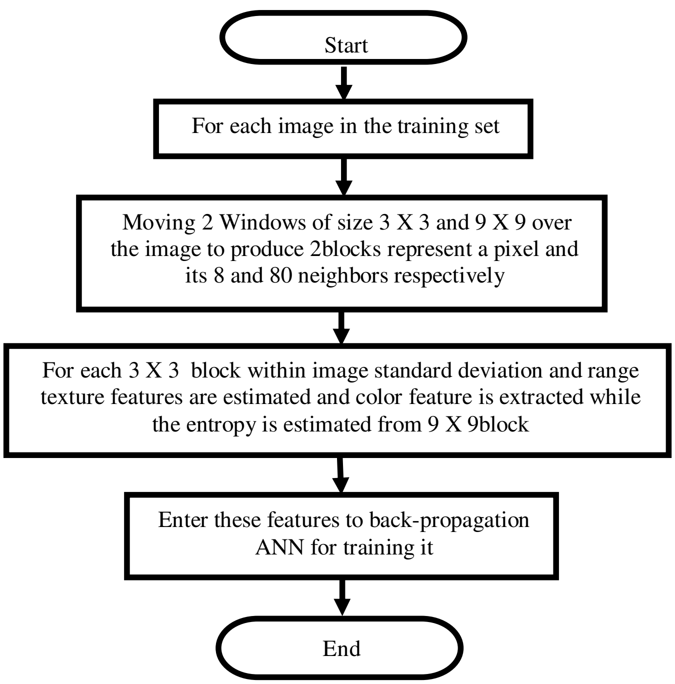

Figure 4. Training phase of the proposed skin detector.

matrices (Red, Green, and Blue), and their average is determined. Figure 3 shows (for example) the scheme of computing entropy for multi-channel image matrices.

In the training phase of the ANN, the weight matrices between the input and the hidden and output layers are initialized with random values. After repeatedly presenting features of the input samples and desired targets, we compare the output with the desired outcome, followed by error measurement and weight adjustment until the correct output for every input is attained. Furthermore, the hidden layer neurons are estimated using an activation function that features the hyperbolic tangent sigmoid transfer function, whereas, the output layer neuron is estimated using the activation function that features the linear transfer function. The training algorithm used is Gradient descent with momentum back propagation. To train the neural network for skin detection, features are extracted and entered as training input data into the ANN. The quality of the training sets that enters into the network determines how well the detector performs. The training phase of the skin detector is illustrated in Figure 4.

ANN testing phase

Throughout this phase, the features are extracted in the same manner as in training phase, then each pixel of a given image (test image) is tested depending on training data. If a pixel is detected as skin, it will be stored in a new image (skin image) at the position of the original image. After examining all image pixels, a new binary image is obtained, which includes only skin pixels. In addition, this phase is used in the performance evaluation of the proposed skin detector.

RESULTS

The experimental results are presented to show the effectiveness of the proposed skin detector, which is based on integrating both texture and color features. Our skin detection system was carried out on a 3.00 GHz Intel (R) Core TM 2Duo processor with 8 GB RAM on Windows Vista platform using MATLAB R2009b. In order to find the overlap between skin color tones and non skin, three histograms have been constructed. A number of

Table 1. Statistical summary gathered from the three histograms.

Color model	Total counts	Total occupied bins	Percentage unoccupied (%)
General color model	67194616	1168154	44.29
Skin model	9244195	211167	89.93
Non-skin model	45450410	1116794	46.74

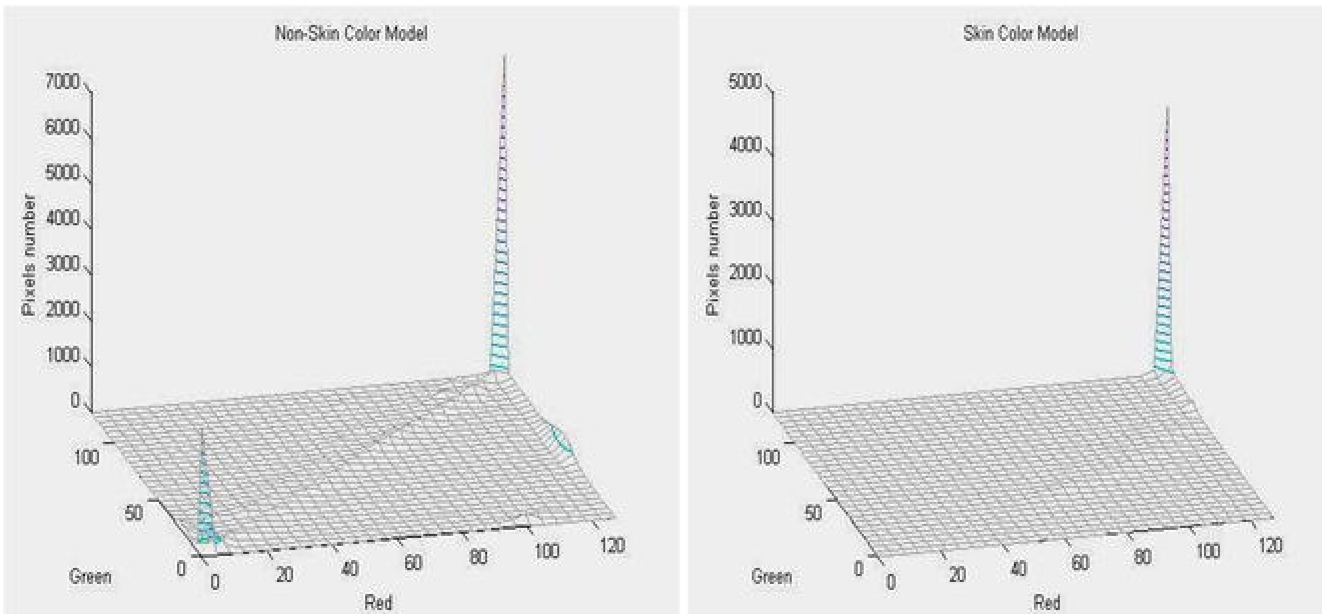

Figure 5. Non-skin and skin color model histograms along the red and green axes.

statistics about the general, skin, and non-skin histogram color models are summarized in Table 1. The total counts field gives the total number of pixels used to form each of the three models. It is worth noting that the skin model is formed by more than (9244195) hand labeled skin pixels. Total occupied bins refers to the number of bins in each model with non- zero counts. This is also expressed as the percentage of the bins in each model that are unoccupied.

The main observation on the statistics shown in Table 1 is the 44.29% of the 2.097 million possible colors in the RGB color, space model are not encountered in any of the training images; 211.167 colors are reflected as skin and 1,116.794 as non-skin. Figure 5 shows the overlap between skin and non-skin colors, suggesting that the skin detection problem based on color features is a difficult issue because a significant overlap exists between the skin and non-skin models. However, overlap is a significant problem only if the counts in the shared bins are comparable in the skin and non-skin cases. The effect of this problem can be minimized using the texture analysis-based features.

Figure 6 shows the sorting distribution of skin versus non-skin pixel statistical features over RGB color. A low rate of overlap among skin, non-skin for standard deviation, and range features was found, whereas, a non low rate of overlap in the entropy feature was observed.

Skin detector testing

The implementation of the skin detector is tested in different images with simple and complex backgrounds, indoor and outdoor settings, as well as different image sizes and skin colors. The experiment is performed on the testing set, which includes 1,119.129 uncontrolled (different illumination, captured quality, distance to camera, etc.) pixels. Each of the first 608.129 pixels belongs to an arbitrary number of skin images and images containing an arbitrary number of people and faces. The other 511,000 pixels reflect no skin pixels, and pixels belonging to images with objects that present skin-like tones (such as a red flower, dog, chocolate, etc.) are included as well.

Three different skin detectors have been tested and evaluated to select the one with higher reliability; then, the evaluations are compared with the performance of previous skin detectors. The first skin detector detects skin pixels with pre-defined rules on skin color tones (Kovac et al., 2003; Gasparini and Schettini, 2006); the

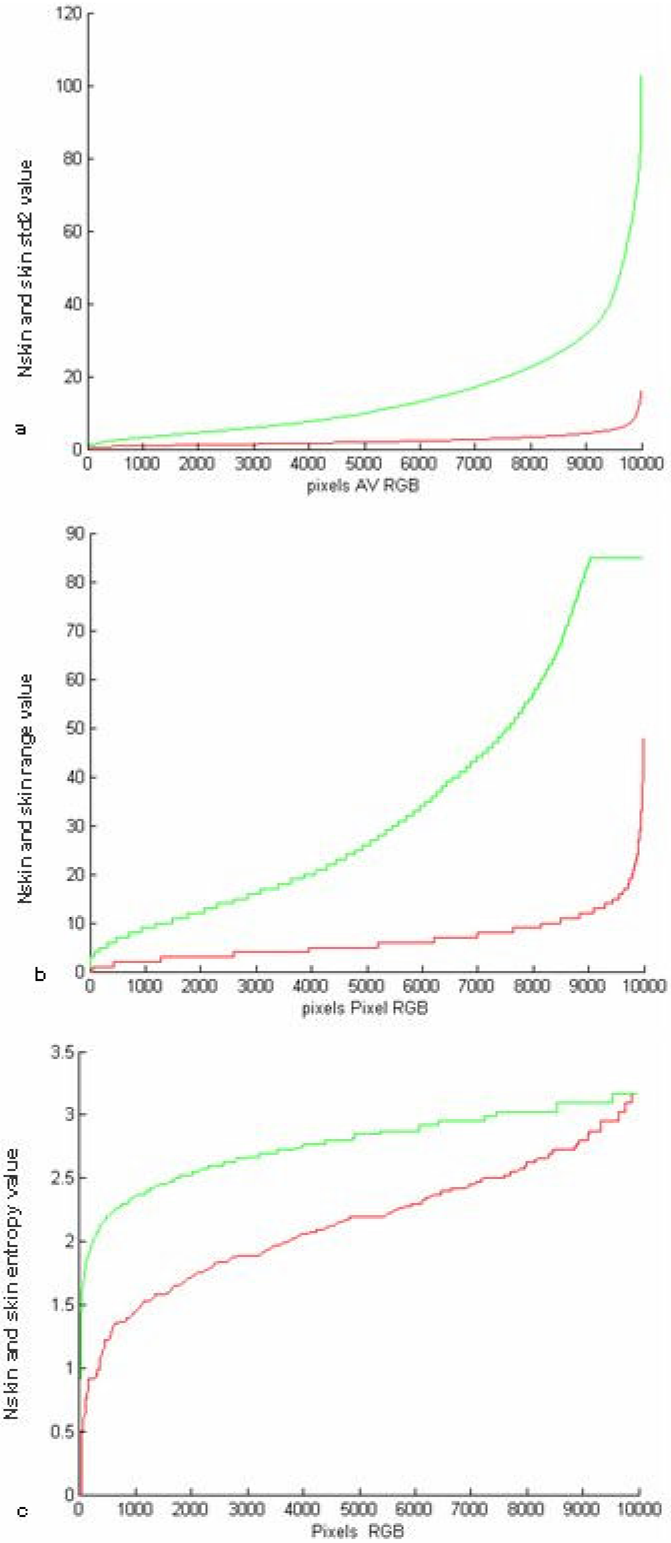

Figure 6. Skin versus non skin intersection texture based features a-for Standard deviation; b-for Range; c- for entropy.

second detects skin pixels based only on texture features; and the last one detects the skin pixels by combining both texture and color features. Figure 7 shows different results obtained by skin pixel detection based on pre-defined skin color rules. The image obtained (Figure 7b) shows that many non-skin pixels are detected incorrectly as skin pixels are comparable with the original image (Figure 7a). The comparison between the skin segment in Figure 6d and the original image in Figure 7c shows that many skin pixels are mistaken for non-skin pixels. Conversely, no skin pixels are detected in the images shown in Figures 7e and f.

The testing result of the skin detector based on ANN with texture features only are shown in Figure 8, including high rates of false positives and false negatives, caused by using only texture features. Many skin regions with different colors noticeably have the same texture; thus, obtaining highly accurate detection rates based only on texture features is impossible.

Figure 9 show that different pixel regions (skin and non-skin) have a variety of colors with similar textures. Some of which (Figures 9a, b, d, and e) have been detected incorrectly using ANN with only texture features.

The testing result of our proposed skin detector based on ANN, combining both texture and color features, is illustrated in Figure 10. The results include high rates of true positives and true negatives with low rates of false detection. Although the image in Figure 9g reflects several human skin types with different colors and textures, the skin pixels within this image are detected correctly by the proposed skin detector, except for a few scattered pixels that is incorrectly detected as non-skin. Most skin pixels within the images shown in Figure 10a are detected correctly (Figure 10b), while no false detection rates are shown within images i and j in Figure 10.

More problematic are images with wood, copper-colored metal, or chocolate colors as they contain shades often occurring in the skin model and are difficult to reliably discriminate using only color or texture features. The combination of both features results in fairly dense sets of true positives; however, the image shown in Figure 10l versus its original (Figure 10k) is correctly classified because it is very similar to skin in color, except for a few scattered pixels incorrectly detected as skin.

THE PROPOSED SKIN DETECTOR EVALUATION

Two types of evaluations are discussed throughout this section. The first evaluation involves the ANN topology, while the second involves the efficiency of the proposed skin detector. Figure 11 shows the performance of the trained network, determined by performing a validation performance analysis between the mean squared error and the number of epochs. The best validation performance is 0.13551 at an epoch of 1000. The regression analysis (Figure 12) returns the correlation coefficient R between the output and the target for training, validation, testing and others.This coefficient equals to 0.928; thus,

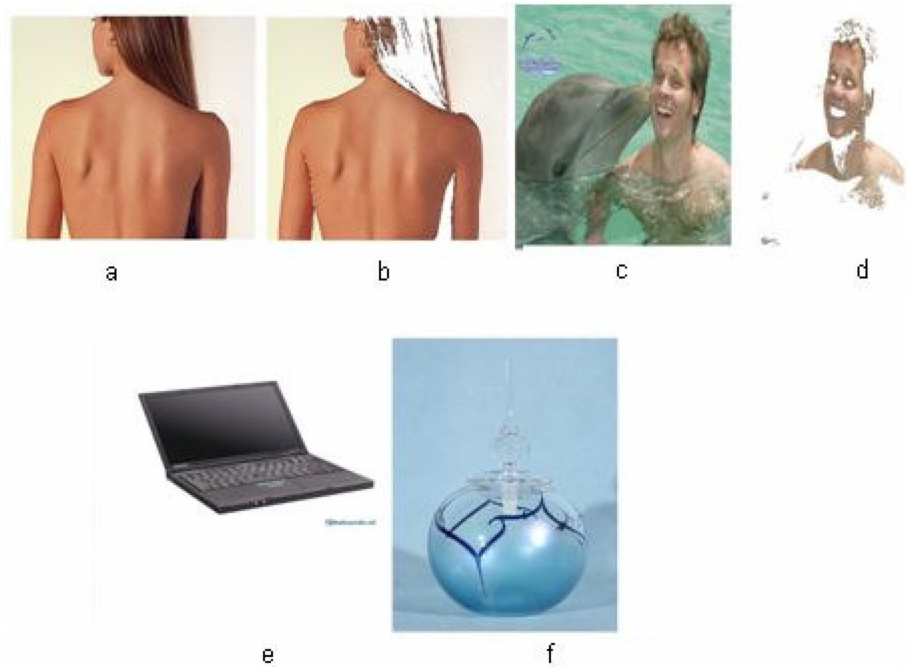

Figure 7. Examples of results obtained by using pre-defined skin color rules based skin detector.

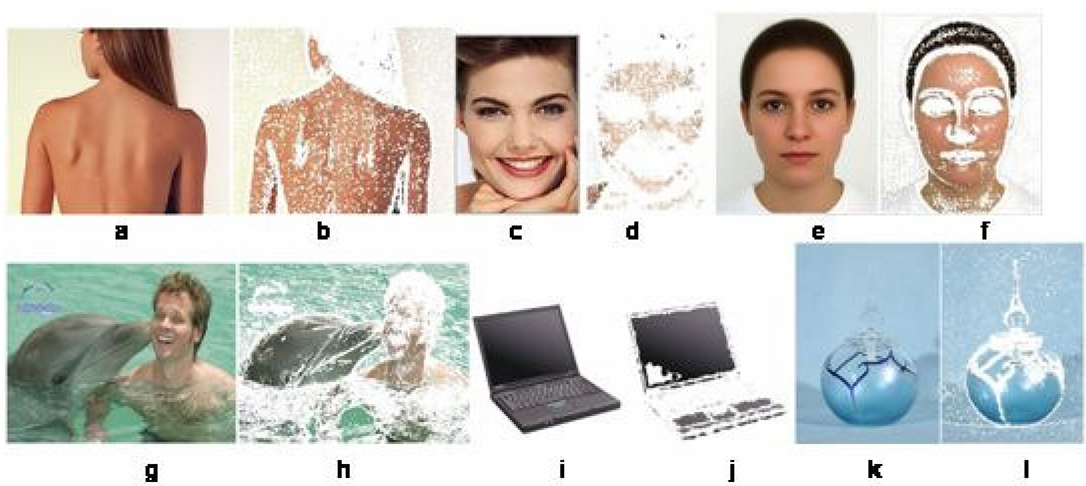

Figure 8. Results of testing skin detector using ANN based on texture features only.

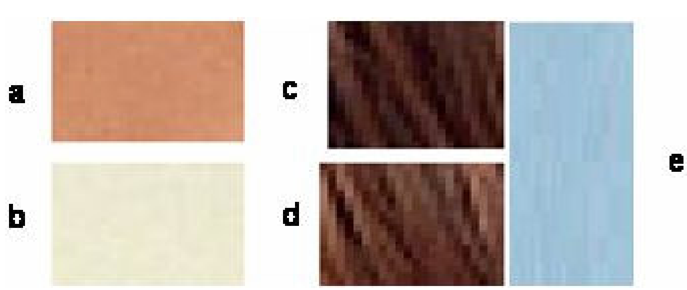

Figure 9. Different pixel regions have variety colors with similar texture.

both output and target are very close, indicating good fit. A skin detection process is never perfect and different users use varying criteria for performance evaluation. General appearance of size zones detected is one of the evaluation criteria. To quantify performance evaluation, True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) are computed for all pixels in the "skin classifier testing set" through skin detector testing. FP is the proportion of non-skin pixels classified incorrectly as skin, whereas, TP is the proportion of skin pixels classified correctly as skin. TN and FN are the

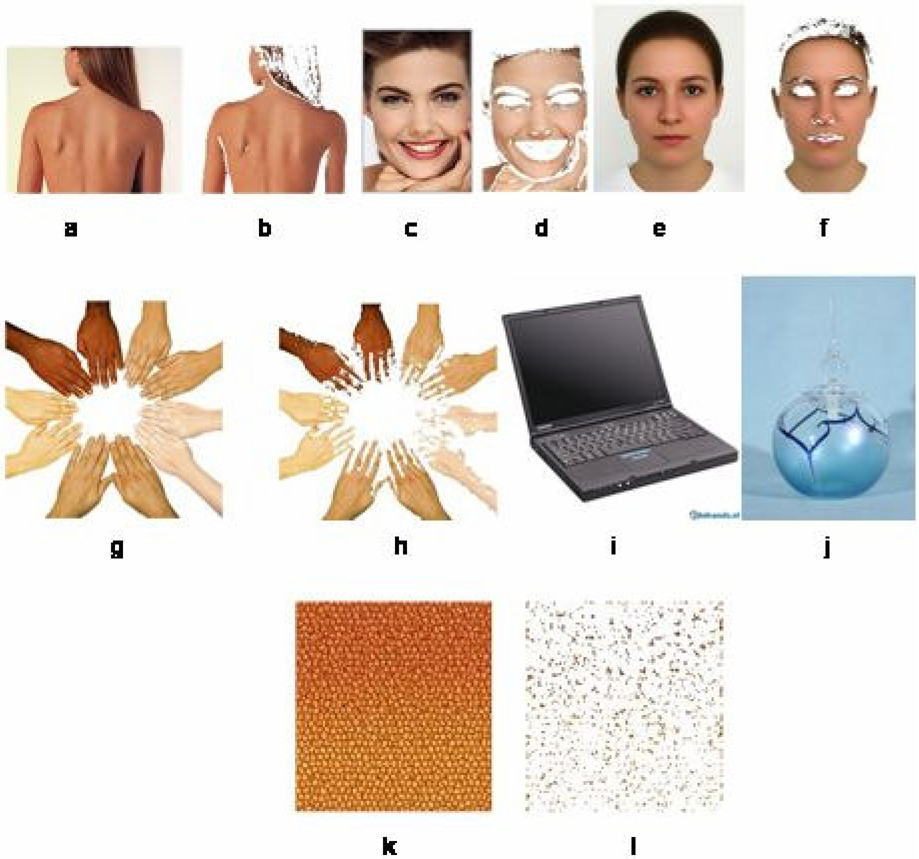

Figure 10. Results of testing our proposed skin detector using ANN based on combining both texture and color features.

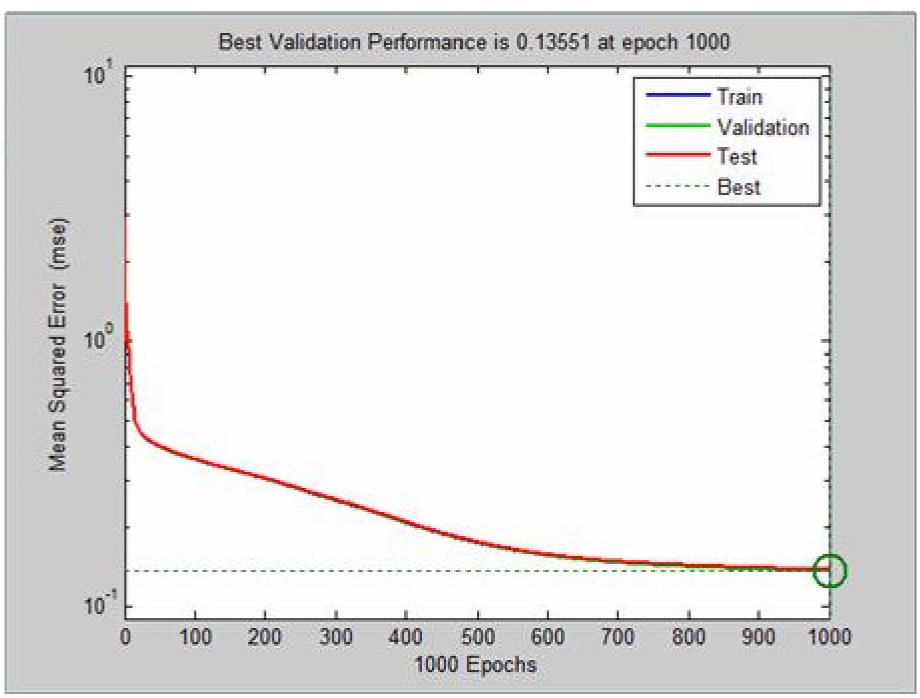

Figure 11. Neural network training performance.

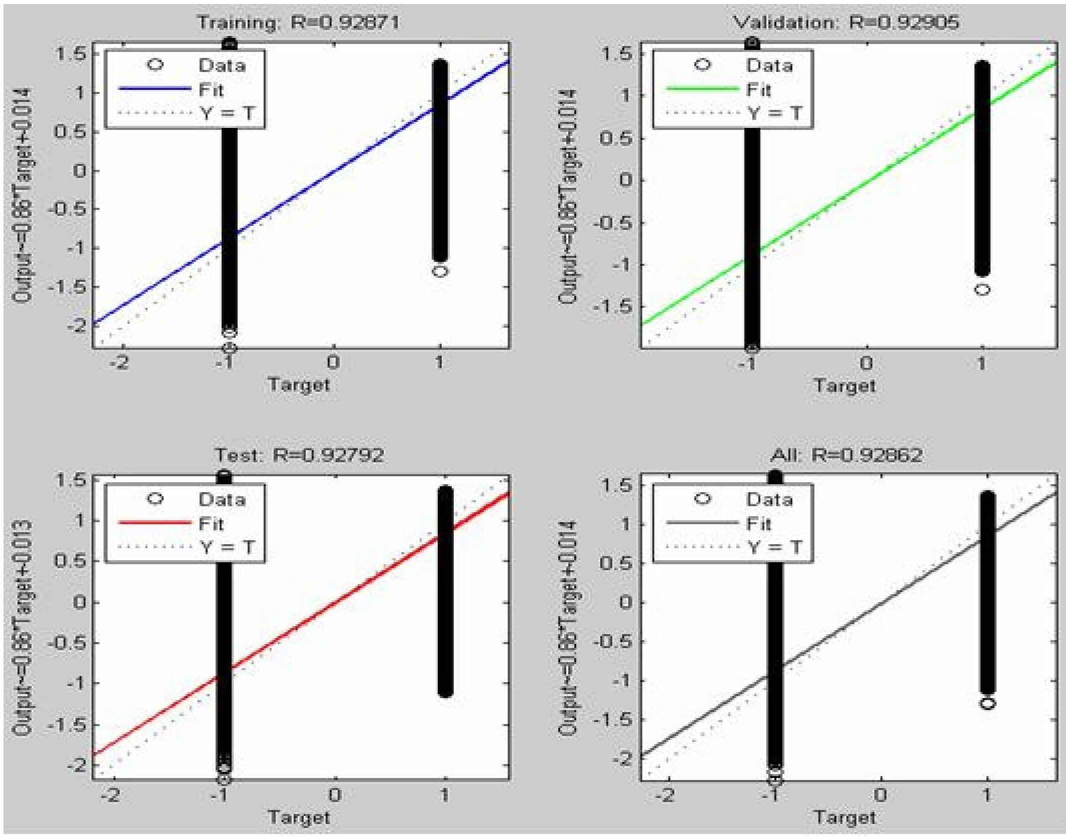

Figure 12. Regressions analysis.

Table 2. True positive and false positive for skin detectors.

Skin detector	TP(%)	FP (%)
Skin detector based on pre-defined color tone rules	98.8891	7.3173
Skin detector based on texture feature only	71.7051	61.9333
Skin detector based on combing both texture and color features	95.6176	0.8795

Table 3. Evaluation metrics for different skin detectors

Skin detector	Recall	Precision	Specificity	Accuracy
Pre-defined	0.988	0.575	0.926	0.932
Texture feature only	0.171	0.105	0.388	0.418
Merge both color and texture features	0.956	0.916	0.991	0.998

complements of FP and TP, respectively. To evaluate the skin detector by three approaches (pre-defined color tone rules, texture features only, and a combination of both color and texture features), a sample of 1,119.129 pixels are used in this evaluation (Table 2).

Four metrics (Table 3) are used to evaluate the performance of the three skin detectors. The first metric is recall or sensitivity, and is computed as (Gasparini et al., 2005; Fawcett, 2003):

Recall =TP/(TP+FN) (5)

The second metric is precision, and is computed as (Gasparini et al., 2005; Fawcett, 2003):

Precision= TP/(TP+FP) (6)

The last two metrics are specificity and accuracy, and are computed respectively as (Gasparini et al., 2005):

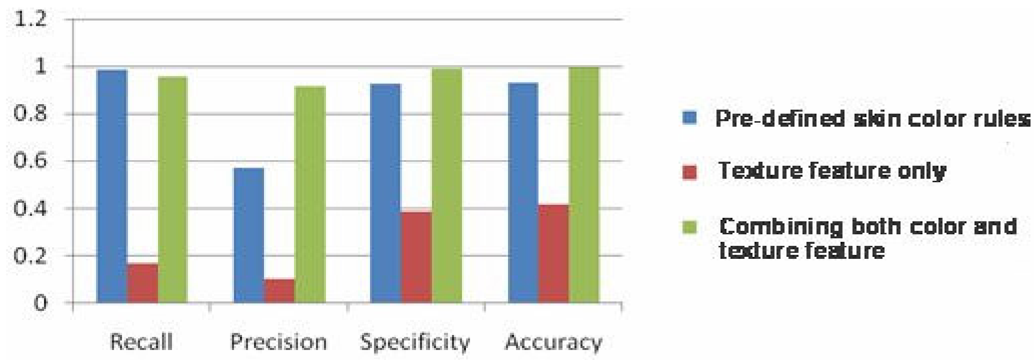

Figure 13. Evaluation metrics for different skin detectors.

Specificity= TN/(FP+TN) (7)

Accuracy = (TP+TN)/(TP+FN+FP+TN) (8)

DISCUSSION

Considering the unconstrained nature of Internet images, the performance of the skin detector is surprisingly good. The best performance can detect 95.6176% of skin pixels with a FP rate of 0.8795%, by combining both texture and color features. A simple comparison among the performance evaluation of pre-defined-based skin detector, texture-based skin detector, and a combination of both color and texture features is shown in Figure 13.

Although there is no means to locate any two papers, which used the same test sets, examining previously published results may be useful. The performance of the proposed skin detector in this research is compared with the performance of other skin detectors. The Bethe Tree Approximation of First Order Model proposed by Zheng et al. (2004) can detect 72% of skin pixels with a 5% FP rate, whereas, the proposed Bayesian model by Jones and Rehg (2000) can detect 69% at the same FPs. Likewise, this model can detect 80% of skin pixels with an 8.5% FP rate, or 90% correct detection with 14.2% FPs. The recall rate of pixel-based skin color classification proposed by Gasparini et al. (2005) is 92%, while precision is equals 39%. The best performance of the skin detector based on the Bayes decision rule Taqa and Al-Sulaifanie (2010) has a detection rate of 94.29% of skin pixels with FPs equaling 6.48%. These evaluation metric values indicate that the proposed skin detector outperforms and is the most reliable of all the aforementioned skin detectors.

CONCLUSION AND FUTURE WORKS

Skin detection is an important pre-process in many image analysis applications; hence, we proposed an improved skin detection method integrating both color and texture features to increase the reliability of skin detection using ANN. The neighborhood information of each pixel was also used through the training and testing phases. More than one approach (pre-defined rules of skin color tones, texture-based skin detector, and a combination of both color and texture features) were applied and tested. We have shown in this paper that a skin detector based on combined color and texture features can lead to an efficient and more reliable method for supervised skin detection compared with using pre-defined color tone rules or texture features only. The proposed detector reduces the FP rate to 0.8795% with respect to pre-defended color tone rules and based on texture feature-only skin detectors. A necessary future direction is to validate the proposed algorithms using a standard skin database data set. Such a method will enable us to compare our detection results with those presented by other authors for the same test images. Another improvement is the adaptation of our approach to Fuzzy logic technique in other domains.

REFERENCES

Ahmed A, Javed Y, Qayyum U (2007). A robust fuzzy logic based approach for skin detection in colored images , International Association Of Science And Technology For Development Proceedings of The Eleventh IASTED International Conference on Artificial Intelligence and Soft Computing, SESSION: Machine vision and intelligent robots pp. 295-301 ,ISBN:978-0-88986-694-2, The publisher ACTA Press Anaheim, CA, USA.

Brown D, Craw I, Lewthwaite J (2001). A SOM Based Approach to Skin Detection with Application in Real Time Systems, In Proceedings of British Machine Vision Conference (BMVC), Manchester,UK,10-13 September.

Chan Y, Harvey R, Smith D (1999). Building Systems to Block Pornography, In 2nd UK Conference on Image Retrieval: The Challenge of Image Retrieval (CIR'99), BCSEI Electronic Workshops in Computing Series, New Castle, UK, pp. 34-40.

Dadgostar F, Sarrafzadeh A, Overmyer SP (2005). Face Tracking Using Mean-Shift Algorithm: A Fuzzy Approach for Boundary Detection, presented at Affective Computing and Intelligent Interaction, Beiging, China.

Fawcett T (2003). ROC Graphs: Notes and practical Consideration for Researchers, Technical Report HPL, Hp Lab, Palo Alto, California,

USA.
Fleck M, Forsyth DA, Bregler C (1996). Finding Naked People, Proc. European Conf. Comput. Vision, 2: 593-602.
Gasparini F, Corchs S, Schettini R (2005). Pixel Based Skin Colour Classification Exploiting Explicit Skin Cluster Definition Methods, In Proceedings of International Colour Association (AIC) Colour, Granda, Spain, pp. 543-546.
Gasparini F, Schettini R (2006). Skin Segmentation using Multiple Thresholding, Internet Image VII Conference, San Jose, California, USA.
Gavrila DM (1999). The analysis of human motion and its application for visual surveillance. Second IEEE Workshop on visual surveillance, pp. 3-5.
Ghouzali S, Hemami S Rziza M, Aboutajdine D, Mouaddib EM (2008). A skin detection algorithm based on discrete cosine transform and generalized Gaussian density, Image processing, ICIP 15th IEEE International Conference, ICIP08 IEEE, pp. 605-608.
Gonzalez RC,Woods RE (2002). Digital Image Processing, 2nd edition, Prentice Hall, Inc.
Jones M, Rehg M (2000). Statistical Color Models with Application to Skin Detecting, Int. J. Comput. Vision (IJCV), 1(46): 81-86.
Kovac J, Peer P, Solina F (2003). Human Skin Colour Clustering for Face Detection, In International Conference on Computer as a Tool, (EUROCON), Ljubljana, Slovenia.
Lin YC, Tseng HW, Fuh CS (2003). Pornography Detection Using Support Vector Machine, 16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP), Kinmen.
Maskooki A, Gunawan E, Soh CB, Low KS (2009). Low frequency Domain Skin Artifact Removal Method for Ultra-Wideband Breast Cancer Detection, Progress Electromagnetic Res., PIER 98: 299-314.
Ruiz-del-Solar J, Verschae R (2004). Robust Skin Segmentation using Neighborhood Information, ICI.
Schalkoff R (1999). Artificial Neural Networks,McGraw-Hill.
Singh SK, Chauhan DS, Vatsa M, Singh R (2003). A Robust Skin Color Based Face Detection Algorithm Tamkang J. Sci. Eng., 6: 4.
Sivanandam SN, Sumathi S, Deepa SN (2006). Introduction to Neural Networks using MATLAB 6.0 Tata McGraw-Hill Publishing.
Subramanian S, Kumar K, Mishra B, Banerjee A, Bhattacharya D (2008). Fuzzy Logic based Content Protection for Image Resizing by Seam Carving, Soft Computing in Industrial Applications, SMCia '08. IEEE Conference on 25-27 June.
Taqa A, AL-Sulaifanie B (2010). Constructing Pornographic Images Detector based on naïve Bayesian classifier, J. Education Sci. College Education University Mosul, 23(1): 84, ISSN 1812-125X.
Taysi N (2010). Application of Neural Network models on analysis of prismatic structures, Sci. Res. Essays, 5(9): 978-989, 4 May, Available online at http://www.academicjournals.org/SRE ISSN 1992-2248.
Umbaugh ES (1998). Computer Vision and Image Processing , Prentice Hall.
Vezhnevets V, Sazonov V, Andreeva A (2003). A survey on pixel-based skin color detection techniques, GRAPHICON03, pp. 85-92.
Zheng H, Liu H, Daoudi M (2004). Blocking Objectionable Images: Adult Images and Harmful Symbols, Proceedings of IEEE International Conference on Multimedia and Expo, Taipei, Taiwan, 27-30 June.